

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2003, 23:56:58 ; Search time 3074 Seconds
(without alignments)
10736.038 Million cell updates/sec

Title: US-10-091-628-1

Perfect score: 1134

Sequence: 1 atgagagcattgtccag.....acatcattcatgcatag 1134

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

ched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: Geneml:*
2: gb_ba:*
3: gb_hcg:*
4: gb_in:*
5: gb_om:*
6: gb_ov:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_ey:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_ets:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_ey:*
39: em_hngo_hum:*
40: em_hngo_mus:*
41: em_hngo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	377	33.2	23618	9	AC079237 Homo sapi
C 2	377	33.2	192263	2	AC093827 Homo sapi
3	375.4	33.1	65266	2	AC099847 Homo sapi
4	320.4	28.3	2263	6	AR033870 Sequence
5	320.4	28.3	2263	6	AR033870 Sequence
6	320.4	28.3	2263	6	IG2744 Sequence 1
7	307.8	27.1	1629	10	CG002028 Cricetus
8	307.8	27.1	1629	10	D87059 House mouse
9	306.4	27.0	1116	4	AB002693 Mus muscu
10	297.8	26.3	1047	6	Z54357 O. cuniculus
11	297.8	26.3	1047	6	AR033871 Sequence
12	297.8	26.3	1047	6	AR033871 Sequence
13	293.2	25.9	4269	9	HSU10417 Homo sapien
14	265	23.4	106811	2	AC120684 Rattus norv
15	261.8	23.1	215210	10	AL713989 Mouse DNA
C 16	217	19.1	65958	9	AC105413 Homo sapi
17	184	16.2	1212	4	OCU131361 Oryctolag
18	183.2	16.2	1663	6	AX401950 Sequence
19	182.6	16.1	1663	10	RAT5BACT Rattus norv
C 20	182.6	16.1	543	11	G51602 SHGC-79180
21	173.8	15.3	1411	10	MMU95132 Mus musculu
22	173.8	15.3	1579	10	U95132 Mus musculu
23	173.8	15.3	1596	10	AB003303 Mouse mRN
24	173.8	15.3	1649	10	BC021154 Mus muscu
25	173.6	15.3	1580	6	AX409529 Sequence
26	173.6	15.3	1580	9	HUMAN7C L21893 Human Na/c
27	138.2	12.2	1686	9	BC012048 Homo sapi
28	138.2	12.2	1714	4	BC019066 Homo sapi
29	133.2	11.7	27247	9	AF002005 Oryctolag
30	118.8	10.5	2020	9	HS1SD8A1 U67650 Human ila
31	118.8	10.5	7979	9	HSBARTPRO Z54350 Homo sapien
32	118.8	10.5	157352	2	AC009687 Homo sapi
C 33	118.8	10.5	170407	2	AC011177 Homo sapi
C 34	118.8	10.5	171132	9	AL161771 Human DNA
C 35	118.8	10.5	172835	2	AL139797 Homo sapi
36	117	10.3	1364	10	AB078635 Mus muscu
37	117	10.3	173363	2	AC114998 Mus muscu
C 38	117	10.3	182230	2	AC110744 Mus muscu
39	115.6	10.2	562	10	AF181258 Mesocric
40	115.4	10.2	1790	9	BC004966 Homo sapi
41	115.4	10.2	4379	9	HSP3 X12458 Human P3 ge
C 42	115.4	10.2	137217	2	HSR211110 Homo sapi
C 43	115.4	10.2	160656	2	AL592072 Homo sapi
C 44	115.4	10.2	219447	9	AL592072 Homo sapi
45	113.8	10.0	1443	10	AF266724 Mus muscu

ALIGNMENTS

```

RESULT 1
AC079237 23618 bp DNA linear PRI 21-FEB-2002
LOCUS AC079237
DEFINITION Homo sapiens BAC clone Rpl1-711J3 from 4, complete sequence.
ACCESSION AC079237
VERSION AC079237.7 GI:18482358
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 23618)
AUTHORS Sulston, J.E. and Watson, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)

```

MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 23618)
 AUTHORS Radionenko, M. and Meyer, R.
 TITLE The sequence of Homo sapiens BAC clone RP11-711J3
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 23618)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (24-AUG-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 23618)
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (03-FEB-2002) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 23618)
 AUTHORS Waterston, R.
 TITLE Direct Submission
 JOURNAL Submitted (21-FEB-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Feb 3, 2002 this sequence version replaced gi:18151062.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@wustl.wustl.edu
 ----- Summary Statistics
 Center Project name: H_NH0711J03

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tateo, M., Catanesi, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the left is RP11-64A1, 2000 bp overlap; the clone sequenced to the right is RP11-397E7, 2000 bp overlap. Actual start of this clone is at base position 1995 of RP11-64A1; actual end is at base position 108789 of RP11-397E7.
 Location/Qualifiers
 1. 23618
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="4"
 /map="4"

/clone="RP11-711J3"
 /clone_id="RP11-11"
 595.720
 /rpt_family="Alu"
 repeat_region 789.1014
 /rpt_family="Alu"
 repeat_region 1039.1072
 /rpt_family="TAAAn"
 repeat_region 1734.1936
 /rpt_family="MIR"
 repeat_region 2029.2175
 /rpt_family="L1"
 2339.2706
 /note="similar to Sus scrofa EST BE031975 (NID:98326984)"
 misc_feature 2484.2656
 /note="match to EST BE181226 (NID:98660402)"
 2499.2656
 /note="similar to Mus musculus EST BB613812 (NID:916454310)"
 misc_feature 2740.2892
 /rpt_family="MIR"
 repeat_region 3617.3920
 /rpt_family="Alu"
 repeat_region 4602.4890
 /rpt_family="Alu"
 repeat_region 4892.4914
 /rpt_family="AT-rich"
 5459.5624
 /rpt_family="Alu"
 repeat_region 6280.6368
 /note="match to EST BE181226 (NID:98660402)"
 6280.6368
 /note="similar to Mus musculus EST BB613812 (NID:916454310)"
 6811.7098
 /rpt_family="Alu"
 repeat_region 7515.7694
 /rpt_family="MIR"
 7794.7912
 /note="similar to Mus musculus EST BB613812 (NID:916454310)"
 7794.7885
 /note="match to EST BE181226 (NID:98660402)"
 8198.8513
 /rpt_family="L1"
 8514.8910
 /rpt_family="MALR"
 8911.9173
 /rpt_family="L1"
 9184.9592
 /rpt_family="MALR"
 9873.10183
 /rpt_family="Alu"
 10157.10266
 /rpt_family="GA-rich"
 10822.11165
 /rpt_family="MALR"
 11242.11261
 /rpt_family="TTTG)n"
 11243.11530
 /rpt_family="Alu"
 11690.11833
 /rpt_family="GA-rich"
 11870.11929
 /rpt_family="L1"
 11916.11938
 /rpt_family="AT-rich"
 12632.12696
 /rpt_family="L1"
 12842.12935
 /rpt_family="L2"
 13129.13424
 /rpt_family="Alu"
 repeat_region

```

repeat_region      13401..13424
                    /rpt_family="(A)n"
repeat_region      13607..13641
                    /rpt_family="L1"
repeat_region      13642..13952
                    /rpt_family="Alu"
repeat_region      13953..14359
                    /rpt_family="L1"
repeat_region      14392..14416
                    /rpt_family="(T)n"
repeat_region      14562..14588
                    /rpt_family="AT-rich"
repeat_region      15001..15392
                    /rpt_family="L1"
repeat_region      16436..16604
                    /rpt_family="MIR"
repeat_region      16895..17214
                    /rpt_family="Alu"
repeat_region      17186..17231
                    /rpt_family="(GAAA)n"
repeat_region      17275..17373
                    /rpt_family="(TTTC)n"
repeat_region      17344..17657
                    /rpt_family="Alu"
repeat_region      17713..17824
                    /rpt_family="L2"
repeat_region      18000..18165
                    /rpt_family="L1"
repeat_region      18235..18384
                    /rpt_family="MIR"
repeat_region      18385..18693
                    /rpt_family="Alu"
repeat_region      18666..18695
                    /rpt_family="AT-rich"
repeat_region      18694..18807
                    /rpt_family="MIR"
repeat_region      18881..19188
                    /rpt_family="MALR"
repeat_region      19189..19562
                    /rpt_family="MALR"
repeat_region      19563..19626

```

Query Match 33.2%; Score 377; DB 9; Length 23618;
 Best Local Similarity 100.0%; Pred. No. 4.9e-104;
 Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1 ATGAGAGCCATTGTTCCAGCAGCTCAGCTCCCTCCCAAGTTTCAGAGAGAGCTG 60
23603 ATGAGAGCCATTGTTCCAGCAGCTCAGCTCCCTCCCAAGTTTCAGAGAGAGCTG 23544

61 CCAGTGGGAGCTGGAGTGATGAGAACTGGAGCTCGTTTTCAGATGGTCCACTGTG 120
23543 CCAGTGGGAGCTGGAGTGATGAGAACTGGAGCTCGTTTTCAGATGGTCCACTGTG 23484

121 ATGATGGGAGCTGGATGTTCTCTTGGAGATGTTCCGAGATCCGAGAGCTGTGTG 180
23483 ATGATGGGAGCTGGATGTTCTCTTGGAGATGTTCCGAGATCCGAGAGCTGTGTG 23444

181 CACATCAGAGAGACCTGGGGAGATTGCTGTGGAGCTGCTCCGAGTTGGGCTCAGCCT 240
23423 CACATCAGAGAGACCTGGGGAGATTGCTGTGGAGCTGCTCCGAGTTGGGCTCAGCCT 23364

241 TTATAGAGCTTATCTCTGGCCATTTAGCTTTTCTGAGAGCCAGTCCAGGCTATGCTGT 300
23363 TTATAGAGCTTATCTCTGGCCATTTAGCTTTTCTGAGAGCCAGTCCAGGCTATGCTGT 23304

301 CTCATCATGAGGCTGTGCTCCGGGGGAGCAACATCTCTAATTTTCACCTTCTGGTTAT 360
23303 CTCATCATGAGGCTGTGCTCCGGGGGAGCAACATCTCTAATTTTCACCTTCTGGTTAT 23244

361 GGAGATATGATCTCAG 377
23243 GGAGATATGATCTCAG 23227

```

```

RESULT 2
AC093827/c
LOCUS
DEFINITION Homo sapiens BAC clone RP11-397E7 from 4, complete sequence.
ACCESSION AC093827 AC016973
VERSION AC093827.3 GI:16328304
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 192263)
Sullivan, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
9063792
MEDLINE
9847074
2 (bases 1 to 192263)
Goya, E., Meyer, R. and Dixon, R.
The sequence of Homo sapiens BAC clone RP11-397E7
Unpublished (2001)
3 (bases 1 to 192263)
Waterston, R.H.
Direct Submission
Submitted (10-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 192263)
Waterston, R.H.
Direct Submission
Submitted (23-OCT-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 192263)
Waterston, R.
Direct Submission
Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 23, 2001 this sequence version replaced gi:15809171.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics
Center project name: H_NH0397E07
Drafting Center: WIBR

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D. McCheserson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
 The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E.,

Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-711J3; the clone sequenced to the right is RP11-168E22. Actual start of this clone is at position 1 of RP11-397E7; actual end is at base position 192263 of RP11-397E7.

Data from AC079237 and AC093779 was used to finish this clone, AC093827. Polymorphisms have been identified between AC079237 and AC093827.

FEATURES

The sequence of AC016973 has been incorporated into AC093827.

```
source
    1..192263
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="4"
        /map="4"
        /clone="RP11-397E7"
        /clone_11b="RPCT-11"
        1..194
            /rpt_family="MIR"
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    3422..3792
        /rpt_family="MALR"
        3917..3942
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    4460..4472
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    4460..4472
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    4750..4772
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    6150..6469
            /rpt_family="MALR"
            6470..6781
                /rpt_family="L1"
                /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    6754..6795
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    6823..7127
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    7245..7324
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    7386..7651
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    7721..7831
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    7853..8154
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    7967..7987
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    8134..8165
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    8269..8410
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    9108..9235
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    9366..10007
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    10729..11025
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    11003..11025
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    11207..11818
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    11808..11850
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
repeat_region
    11851..12151
            /rpt_family="L1"
            /note="similar to EST BE151388 (NID:g8614109)"
```

```
misc_feature
    12109..12705
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    12132..12151
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    12788..12884
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    13357..13649
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    13735..14280
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    14376..14499
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    14500..15875
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    15857..15880
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    15876..16047
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    16086..16231
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    16254..16690
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    21991..22029
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    22069..22413
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    22460..22610
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23018..23063
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23034..23552
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23040..23737
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23042..23907
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23042..23852
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23042..23791
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23042..23696
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23042..23491
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23043..23763
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23043..23273
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23044..23864
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23053..23874
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23055..23873
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23056..23907
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
repeat_region
    23056..23907
        /rpt_family="L1"
        /note="similar to EST BE619594 (NID:g13670965)"
```

Query Match 33.2%; Score 377; DB 9; Length 192263;
Best Local Similarity 100.0%; Pred. No. 6.9e-104; Mismatches 0; Indels 0; Gaps 0;
Matches 377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCAAATGTTCCAGAGCTCAGCTGCTCCGCAACAGTTCCAGAGAGAGCTG 60
DB 1985 ATGAGAGCCAAATGTTCCAGAGCTCAGCTGCTCCGCAACAGTTCCAGAGAGAGCTG 1926
QY 61 CCAGTGGAGCTGAGAGCTGAGAGAACTGAGAGCTGTTTCCAGCTGCTGCTG 120
DB 1925 CCAGTGGAGCTGAGAGCTGAGAGAACTGAGAGCTGTTTCCAGCTGCTGCTG 1866
QY 121 ATGATGGAGCTGCTGATGCTCTTGGAGATGTTCCGAGAGATCCGAGAGCTGCTG 180

Db 1865 ATGATGGGCGTGTCTATGTTCTTTGGAGTTCCTGGAGATCGGAAGCTGTGTCG 1806
Oy 181 CACATCAGAGAGACCCCTGGGGGAGTGTGTGGAGCTGCTGTGCGAGTTGGGCTAGTGCCT 240
Db 1805 CACATCAGAGAGACCCCTGGGGGAGTGTGTGGAGCTGCTGTGCGAGTTGGGCTAGTGCCT 1746
Oy 241 TTACAGCTTTATCTCTCTGCGCATTTAGCTTTTCTGGAAGCCAGTCCAGCTATTGCTGTT 300
Db 1745 TTACAGCTTTATCTCTCTGCGCATTTAGCTTTTCTGGAAGCCAGTCCAGCTATTGCTGTT 1686
Oy 301 CTCATCATGGGCTGCTGCCCGGGGGGACCATCTCTAACATTTTACCTTTGGGTTGAT 360
Db 1685 CTCATCATGGGCTGCTGCCCGGGGGGACCATCTCTAACATTTTACCTTTGGGTTGAT 1626
Oy 361 GGAGATATGATCTCAG 377
Db 1625 GGAGATATGATCTCAG 1609
LOCUS 1865
DEFINITION Homo sapiens chromosome 18 clone RP11-819K4 map 18, LOW-PASS
ACCESSION AC099847
VERSION AC099847.1 GI:17047210
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 65268)
Birtten,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 18, clone RP11-819K4
Unpublished
2 (bases 1 to 65268)
Birtten,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barra,N., Bastien,V., Boguslavsky,L., Bouckhalter,B.,
Brown,A., Camarata,J., Campioano,A., Chang,J., Chazaro,B.,
Chepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fairo,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Gilde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Iarocque,K.,
Lamatzes,R., Landers,T., Lehotzky,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mienna,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Rette,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,N., Schupbach,R.,
Seman,S., Severy,P., Spencer,B., Stange-Thomann,S., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.T., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
TITLE Whitehead Institute/MIT Center for Genome Research
JOURNAL Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L13211
Center clone name: 819_K_4

* NOTE: This record contains 81 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1
* 728 827: contig of 727 bp in length
* 828 1547: contig of 720 bp in length
* 1548 1647: gap of 100 bp
* 1648 2369: contig of 722 bp in length
* 2370 2469: gap of 100 bp
* 2470 3148: contig of 679 bp in length
* 3149 3248: gap of 100 bp
* 3249 3922: contig of 674 bp in length
* 3923 4022: gap of 100 bp
* 4023 4708: contig of 686 bp in length
* 4709 4808: gap of 100 bp
* 4809 5512: contig of 704 bp in length
* 5513 5612: gap of 100 bp
* 5613 6320: contig of 708 bp in length
* 6321 6420: gap of 100 bp
* 6421 7119: contig of 699 bp in length
* 7120 7219: gap of 100 bp
* 7220 7943: contig of 724 bp in length
* 7944 8043: gap of 100 bp
* 8044 8780: contig of 737 bp in length
* 8781 8880: gap of 100 bp
* 8881 9602: contig of 722 bp in length
* 9603 9702: gap of 100 bp
* 9703 10429: contig of 727 bp in length
* 10430 10529: gap of 100 bp
* 10530 11235: contig of 706 bp in length
* 11236 11335: gap of 100 bp
* 11336 12030: contig of 695 bp in length
* 12031 12130: gap of 100 bp
* 12131 12822: contig of 692 bp in length
* 12823 12922: gap of 100 bp
* 12923 13644: contig of 722 bp in length
* 13645 13744: gap of 100 bp
* 13745 14466: contig of 722 bp in length
* 14467 14566: gap of 100 bp
* 14567 15297: contig of 731 bp in length
* 15298 15397: gap of 100 bp
* 15398 16109: contig of 712 bp in length
* 16110 16209: gap of 100 bp
* 16210 16935: contig of 726 bp in length
* 16936 17035: gap of 100 bp
* 17036 17749: contig of 714 bp in length
* 17750 17849: gap of 100 bp
* 17850 18542: contig of 693 bp in length
* 18543 18642: gap of 100 bp
* 18643 19341: contig of 699 bp in length
* 19342 19441: gap of 100 bp
* 19442 20144: contig of 703 bp in length
* 20145 20244: gap of 100 bp
* 20245 20963: contig of 719 bp in length
* 20964 21063: gap of 100 bp
* 21064 21779: contig of 716 bp in length
* 21780 21879: gap of 100 bp
* 21880 22596: contig of 717 bp in length
* 22597 22696: gap of 100 bp
* 22697 23361: contig of 665 bp in length
* 23362 23461: gap of 100 bp
* 23462 24186: contig of 725 bp in length
* 24187 24286: gap of 100 bp
* 24287 25016: contig of 730 bp in length
* 25017 25116: gap of 100 bp

25117 25835: contig of 719 bp in length
25836 25935: gap of 100 bp
25936 26607: contig of 672 bp in length
26608 26707: gap of 100 bp
26708 27297: contig of 590 bp in length
27298 27397: gap of 100 bp
27398 28099: contig of 702 bp in length
28100 28199: gap of 100 bp
28200 28896: contig of 697 bp in length
28897 28996: gap of 100 bp
28997 29707: contig of 711 bp in length
29708 29807: gap of 100 bp
29808 30519: contig of 712 bp in length
30520 30619: gap of 100 bp
30620 31348: contig of 729 bp in length
31349 31448: gap of 100 bp
31449 32177: contig of 729 bp in length
32178 32277: gap of 100 bp
32278 33007: contig of 730 bp in length
33008 33107: gap of 100 bp
33108 33808: contig of 701 bp in length
33809 33908: gap of 100 bp
33909 34618: contig of 710 bp in length
34619 34718: gap of 100 bp
34719 35387: contig of 669 bp in length
35388 35487: gap of 100 bp
35488 36158: contig of 671 bp in length
36159 36258: gap of 100 bp
36259 36957: contig of 699 bp in length
36958 37057: gap of 100 bp
37058 37764: contig of 707 bp in length
37765 37864: gap of 100 bp
37865 38561: contig of 697 bp in length
38562 38661: gap of 100 bp
38662 39375: contig of 714 bp in length
39376 39475: gap of 100 bp
39476 40191: contig of 716 bp in length
40192 40291: gap of 100 bp
40292 41017: contig of 726 bp in length
41018 41117: gap of 100 bp
41118 41764: contig of 647 bp in length
41765 41864: gap of 100 bp
41865 42593: contig of 729 bp in length
42594 42693: gap of 100 bp
42694 43356: contig of 663 bp in length
43357 43456: gap of 100 bp
43457 44158: contig of 702 bp in length
44159 44258: gap of 100 bp
44259 44951: contig of 693 bp in length
44952 45051: gap of 100 bp
45052 45766: contig of 715 bp in length
45767 45866: gap of 100 bp
45867 46581: contig of 715 bp in length
46582 46681: gap of 100 bp
46682 47400: contig of 719 bp in length
47401 47500: gap of 100 bp
47501 48234: contig of 734 bp in length
48235 48334: gap of 100 bp
48335 49057: contig of 723 bp in length
49058 49157: gap of 100 bp
49158 49877: contig of 720 bp in length
49878 49977: gap of 100 bp
49978 50699: contig of 722 bp in length
50700 50799: gap of 100 bp
50800 51494: contig of 695 bp in length
51495 51594: gap of 100 bp
51595 52298: contig of 704 bp in length
52299 52398: gap of 100 bp
52399 53096: contig of 698 bp in length
53097 53196: gap of 100 bp
53197 53903: contig of 707 bp in length
53904 54003: gap of 100 bp
54004 54708: contig of 705 bp in length

* 54709 54808: gap of 100 bp
* 54809 55524: contig of 716 bp in length

Query Match
Best Local Similarity 33.1%; Score 375.4; DB 2; Length 65268;
Matches 376; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAGGCCAATGTTCCAGCAGCTCAGCTCCCTGCCACAGTTCAAGAGAGACTG 60
50030 ATGAGAGCCCATGTTTCCAGCAGCTCAGCTCCCTGCCACAGTTCAAGAGAGACTG 50089

61 CCAGTGGAGCTGAGAGTGTAGAGAACTGAGAGCTGTTTACAGTGTGCTGCTG 120
50090 CCAGTGGAGCTGAGAGTGTAGAGAACTGAGAGCTGTTTACAGTGTGCTGCTG 50149

121 ATGATGGGCTGCTCATGTTCTCTTGGAGATGTTCCGTGAGATCCGAGAGCTGCTG 180
50150 ATGATGGGCTGCTCATGTTCTCTTGGAGATGTTCCGTGAGATCCGAGAGCTGCTG 50209

181 CACATCAGAGACCTCGGGGCAATTCCTGTGGAGCTGCTTGCACGTTGGGCTCATGCT 240
50210 CACATCAGAGACCTCGGGGCAATTCCTGTGGAGCTGCTTGCACGTTGGGCTCATGCT 50269

241 TTTACAGCTTATCTCTGCTGCTGCTTCTCTGAGCCAGTCCAGCTATGCTGT 300
50270 TTTACAGCTTATCTCTGCTGCTGCTTCTCTGAGCCAGTCCAGCTATGCTGT 50329

301 CTCATCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
50330 CTCATCATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 50389

361 GGAGATATGATCTCAG 377

50390 GGAGATATGATCTCAG 50406

RESULT 4
AR033870
LOCUS AR033870 2263 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5869265.
ACCESSION AR033870
VERSION AR033870.1 GI:5949475
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2263)
AUTHORS Dawson,P.A.
TITLE Illeal bile acid transporter compositions and methods
JOURNAL Patent: US 5869265-A 1 09-FEB-1999;
FEATURES
source
1..2263
/organism="unknown"

BASE COUNT 672 a 451 c 476 g 664 t

ORIGIN

Query Match 28.3%; Score 320.4; DB 6; Length 2263;
Best Local Similarity 60.8%; Pred. No. 9e-87;
Matches 522; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

80 ATGGAACCTGAGCTGTTTTCACAGTGTGTCACGTGATAGTGGGGCTCATGT 139
188 ACGCCATCTTAGCTGTGATGAGACACCTGCTACATCTCTTACCTTGTGATGT 247

140 TCTCTTGGAGATGTTCCGTGAGATCCGAGAGCTGTGCTCACAATCAGAGACCTGGG 199
248 TTTTCATGAGGAGTGTGATGAGTCACTCACAAGTTTCTGGAGACCTAAGGGGCAATGG 307

200 GCATTCGTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 259
308 GCATTCGTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 367

260 CCATTAGCTTTTCTGTGAAGCAGTCCAGATGATTTGCTGTATCATATGAGGCTGCTGCC 319

[illegible]

Db	188	ACGCCATCTCAGGCTGGTATATGACACCGTGTCTCAAAATCTCTAGCCCTTGATATGT	247
Qy	140	TCTCTTTGGAGATGTTCCGTGAGATTCGGAAGCTGTGTGCGACATCAGAGAACCTGTGG	199
Db	248	TTTCCATGAGGGGTGCAATGTGGAACTCCACAAAGTTCTTGAGGACACTAAGGGCGGCATGGG	307
Qy	200	GCATTGCTGTGGACCTGCTCTGCGCAAGTTTGGGCTCATAGCCTTTTACAGCTTATCTCTGG	259
Db	308	GCATCTGTGTGGGCTTCTCTGTGTAGTTTGGAAATCAAGCCCTCTCAAGGTTTCGTCTGT	367
Qy	260	CCATTAGCTTTCTCTGAGGCACTCCAGGATATGTCGTTCTCATATAGGGCGTGTGCC	319
Db	368	CCGTGGCTTTTGGCATCTCTCCAGTGGACGTGTGTGTGTCTATCCAGGAGTTGTGCC	427
Qy	320	CGGGGGGCAACATCTTAAACATTTTACCTTCTGGGTGATGAGATATGATCTCACA	379
Db	428	CTGGAGGAATGCGCTCCAAATATCTAGCCTATTTGGTATGATGGGACATGGAACCTCAGCG	487
Qy	380	TCAGTATGACAACTGTTTCCACCGGTGGCCGCTTGGGAATGATGCCACTCTGCATTTATC	439
Db	488	TTAGATGATACCACTGTCTCCACGCTGTGGCTTGGAAATGATGCCCTTGTGCTCTCA	547
Qy	440	TCTACACCTGTGCTCGAGCTTTCAGACAGATCTCACATCTCTTATCAGAACATAGGA	499
Db	548	TCATATACCAAGATGTGGTTGACTCAGAGAGCATTTGATATCTTATGACAGCATGGCA	607
Qy	500	TTACCTTTGTGTGCTGTGACATTCCTGTGTGCTTGTGTGTATGTAATTTACATGATGC	559
Db	608	CTTCTGTGTGTGTCTTGTATTTCTCTGTTCATTTGGAATGTATGTATACAAATGCG	667
Qy	560	CAAAACAATCCAAATATCTCAAGATTTGGGGCCGTGTGTGTGGGGTCTCTCTGTG	619
Db	668	CCCAAAAAGCAAAATCAATCTTAAATTTGGATCCATGCGAGGTGCAATTCCTCATTTGTC	727
Qy	620	TGTGTGCAAGTGTGTGTGTGTGTCTGTGGGAAAGATCTTGGAAATTCAGACATCACCTTC	679
Db	728	TCATGTGTGTGTGTGGAGAAATCTGTATCCAAATGTGCTGTGACCATTTGAACCCAAAGCTGT	787
Qy	680	TGACCATCAGTTTATCTTCTTTTGAATGGGCATGTCACGGGTTTTCTGTGGCACTTT	739
Db	788	GGATTTATGGAACATATATCTTATCTAGTGTGCTACGGCTGTGGGTTTTTCTGTGCTGAA	847
Qy	740	TTACCCACAGCTTTGGCAAAAGTGCAGAGCAATTTCTTAAAGAACTGAGCTCAGATA	799
Db	848	TTGTGTGTCAACCTGTGTACAGGTGCGGAACAGATTTCTTGGAAACGGGGTTGCAGAA	907
Qy	800	TTGAGATGTGATCAACATGTCTCAAGTTATTTTTCATCTGTGAGACATTTGTCCAGATGT	859
Db	908	CTCAGCTGTGTTCACCAATTTGTGAGCTTTCTTTCACGCCCTGAGAGACCTCAACTTGTGT	967
Qy	860	TGAGTTTCCACTGTGGCCTATGAGACTTTTCCAGCTGATAGATGATTTCTTATTTGTTCAG	919
Db	968	TCACCTTCCCTCATCTACAGCATTTTCCAGATGCGCTTTGCGACATATCTATTAAGAG	1027
Qy	920	CATATCAGACGTACAGAGA	937
Db	1028	CTTATGTGCAATACAGA	1045
RESULT 6			
CGU02028			
LOCUS			
DEFINITION			
Cricetulus griseus Na+ dependent ileal bile acid transporter mRNA,			
complete cds.			
ACCESSION			
U02028			
VERSION			
U02028.1			
GI:455032			
SOURCE			
ORGANISM			
Chinese hamster.			
Cricetulus griseus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;			
Cricetulus.			

Query Match 27.1%; Score 307.8; DB 10; Length 974;
 Best Local Similarity 59.4%; Pred. No. 5.9e-83;
 Matches 522; Conservative 0; Mismatches 357; Indels 0; Gaps 0;

80 ATGGAACCTGAGCTGCTTTTCAAGTGTGTCACATGATGATGAGGCTGCTCATGT 139
 Db ATGCAATTCATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 115
 56 ATGCAATTCATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 115
 140 TCTCTTTGGAGTGTTCCTGAGATGCGAAGCTGTGTCGACATCAGAGACCTTGGG 199
 Db TTTCTATGGGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 175
 116 TTTCTATGGGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 175
 200 GCATTGCTGTGGAGTGTGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 259
 Db GTATCTTGTGGAGTGTTCCTGAGATGCGAAGCTGTGTCGACATCAGAGACCTTGGG 235
 176 GTATCTTGTGGAGTGTTCCTGAGATGCGAAGCTGTGTCGACATCAGAGACCTTGGG 235
 260 CCATTAGCTTTTCTCTGAAGCTGCAAGCTATGCTGTTCTCATCATGAGGCTGCTGCC 319
 236 CTGAGGCTCTGGCATCTTCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAG 295
 320 CGGGGGGACATCTCTTAACATTTTACCTTGTGAGTGAAGATGAGATGAGATGAGATG 379
 Db CTGAGGAGACTGGCTCCATATCTGAGCTATGATGATGATGATGATGATGATGATGATG 355
 296 CTGAGGAGACTGGCTCCATATCTGAGCTATGATGATGATGATGATGATGATGATGATG 355
 380 TCAGATGACAACTGTCACCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 439
 Db TTAGATGACAACTGTCACCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 415
 356 TTAGATGACAACTGTCACCTGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 415
 440 TCTACACCTGCTCTGAGTCTTTCAGAGAACTTCACATTCCTTATCAGAACATAGAA 499
 Db TCTACACCAAGATGAGGCTTACCTGAGGAGAGATGATGATGATGATGATGATGATGATG 475
 416 TCTACACCAAGATGAGGCTTACCTGAGGAGAGATGATGATGATGATGATGATGATGATG 475
 500 TTACCTTGTGCTGCTGACCAATCTCTGAGGCTTGTGCTGATGATGATGATGATGATG 559
 Db TTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
 476 TTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 535
 560 CAAACCAATCCAAATCATCTTCAGATGTCAGAGTGGGCGGCTGTTGGGCGGCTGCTG 619
 Db CACAAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 595
 536 CACAAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 595
 620 TGTGTCAGATGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTCGTGTC 679
 Db TGTATGCTGTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 655
 596 TGTATGCTGTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 655
 680 TGACCATGATTTATCTTTCTTGTGATGTCAGATGTCAGATGTCAGATGTCAGATGTC 739
 656 GGATTAATAGGAACAATATTTCTTATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 715
 740 TTACCCACAGTCTTGGGAG 799
 Db TAGCTGTACACCTGTCAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 775
 716 TAGCTGTACACCTGTCAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 775
 800 TTCAGATGTCATGATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 859
 Db CTCAGCTGTGCTGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 835
 776 CTCAGCTGTGCTGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 835
 860 TGAAGTTCCTGCTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 919
 Db TCACCTTCCCACTCATATATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 895
 836 TCACCTTCCCACTCATATATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 895
 920 CATATCAGATGTCAG 958
 Db TTTATGTCATCAG 934
 896 TTTATGTCATCAG 934

RESULT 8
 AB002693 1629 bp mRNA linear ROD 11-AUG-1995
 LOCUS AB002693
 DEFINITION Mus musculus mRNA for ISBT, complete cds.
 ACCESSION AB002693
 VERSION AB002693.1 GI:1944178
 KEYWORDS ISBT.

SOURCE
 ORGANISM
 Mus musculus (strain:ICR) 8 weeks male cDNA to mRNA.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

1. 1629
 Submitted (07-APR-1997) Tohru Saeki, Kyoto Prefectural University,
 Department of Biological Resource Chemistry, Nakaragi, Shimogamo,
 Sakyo-ku, Kyoto, Kyoto 606, Japan (E-mail:tsaeki@ids.kpu.ac.jp,
 Tel:81-75-703-5663, Fax:81-75-703-5661)
 Location/Qualifiers
 1. 1629
 /organism="Mus musculus"
 /strain="ICR"
 /db_xref="taxon:10090"
 /sex="male"
 /dev_stage="8 weeks"
 50. 1096
 /codon_start=1
 /product="ISBT"
 /protein_id="BAAL9606.1"
 /db_xref="GI:1944179"
 /translation="MDNSVCPNATVCEGDSVPESENALINTVMSVTLITLLM
 VMESGNGVNEHKEFLGHRPMGIFVEFLQFGLPLGFLISVAGSLPQAVVLI
 MGCPGCGNSNIIAWIDGMDLSVSTKSTSLIALGMLCLPVRKMYDSEGLIY
 PYDSIGSLVALVIVPVSFGMFVNKMPQAKIILKISITGTVILIVLIVAGLILYS
 AWLEPLWIIIGTPIRAGSLGFLRLAGQWYRRTVALLETGMONTDLCSITVOL
 SFSPEENLVFTPLIYTVFQVFAVILIGIYVRYRCYGNDAEFLKTDNEMDSR
 SPSDETNKQFQDEK"

BASE COUNT 451 a 315 c 341 g 522 t

ORIGIN

Query Match 27.1%; Score 307.8; DB 10; Length 1629;
 Best Local Similarity 59.4%; Pred. No. 6.4e-83;
 Matches 522; Conservative 0; Mismatches 357; Indels 0; Gaps 0;

80 ATGGAACCTGAGCTGCTTTTCAAGTGTGTCACATGATGATGATGATGATGATGATGATG 139
 Db ATGCAATTCATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 188
 129 ATGCAATTCATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 188
 140 TCTCTTTGGAGTGTTCCTGAGATGCGAAGCTGTGTCGACATCAGAGACCTTGGG 199
 Db TTTCTATGGGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 248
 189 TTTCTATGGGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 248
 200 GCATTGCTGTGGAGTGTGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATG 259
 Db GTATCTTGTGGAGTGTTCCTGAGATGCGAAGCTGTGTCGACATCAGAGACCTTGGG 308
 249 GTATCTTGTGGAGTGTTCCTGAGATGCGAAGCTGTGTCGACATCAGAGACCTTGGG 308
 260 CCATTAGCTTTTCTCTGAAGCTGCAAGCTATGCTGTTCTATCATATGAGGCTGCTGCC 319
 Db CTGAGGCTCTGGCATCTTCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTG 368
 309 CTGAGGCTCTGGCATCTTCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTGTAAGGCTG 368
 320 CGGGGGGACATCTCTTAACATTTTACCTTGTGAGTGAAGATGATGATGATGATGATGATG 379
 Db CTGAGGAGACTGGCTCCATATCTGAGCTATGATGATGATGATGATGATGATGATGATGATG 428
 369 CTGAGGAGACTGGCTCCATATCTGAGCTATGATGATGATGATGATGATGATGATGATGATG 428
 380 TCAGATGACAACTGTCACCTGTCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAT 439
 Db TTAGATGACAACTGTCACCTGTCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAT 488
 429 TTAGATGACAACTGTCACCTGTCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAGTGCAGAT 488
 440 TCTACACCTGCTCTGAGATGTCAGAGATGTCACATTCCTTATCAGAACATAGAA 499
 Db TCTACACCAAGATGAGGCTTACCTGAGGAGAGATGATGATGATGATGATGATGATGATGATG 548
 489 TCTACACCAAGATGAGGCTTACCTGAGGAGAGATGATGATGATGATGATGATGATGATGATG 548
 500 TTACCTTGTGCTGCTGACATCTCTGAGGCTTGTGTCATGATGATGATGATGATGATGATG 559

DB 549 TTCTCTGCTGCTCTCTGTTATCCCTGTTCCCTTGGAATGTTTGAATCACAATGCG 608
QY 560 CAAACCAATCCAAATCATTTCTCAAGATTGGGGCCGTTGGTGGGGCTCCCTCTGG 619
DB 609 CACAAAGAGCAAGATTATATCTTAATGATTCATCAAGGTGAATTTCAATTTGTC 668
QY 620 TGGTCGCACTTGTGTGTGTCTCTGGCGAAGAGATCTTGAATTCAGATCACCTTC 679
DB 669 TCATGCTGTGATTGAGGAATACCTGACCAAGTCCCTGGATCATGAAACCAACTGT 728
QY 680 TGAACCATCACTTCACTTTCTTTGATTGGCCATGTCAAGGTTTCTGCTGGCACTT 739
DB 729 GGAATTATAGGAACAAATATTCCTATAGCTGCTACAGCCGCTTTCTTCTGGCTAGAC 788
QY 740 TTACCCACAGCTTGGCAAGGTGAGGCAATTTCTTGAATCTGAGGCTCAGAAAT 799
DB 789 TAGCTGTCAACCTGTGTACAGGTGCGCAACGATGCTTGGAACTGGAAATGACAGACA 848
QY 800 TTCAGATGTCAATCACCATGCTCCAGTTATCTTTCACTGCTGAGCACTTGTCCAGATGT 859
DB 849 CTCAGCTGTGCTCCACATTTGTACAGCTCTCTCTTCCCCGAGATCTCAACTGTGT 908
QY 860 TGAATTTCCCACTGCTGTGTGACTCTTCCAGCTGATAGATGATTTCTTATTTTGGAG 919
DB 909 TCACCTTCCACATCTATCTATCTGTTTCCAGCTCTCTTGTGACGATGATATATAGGA 968
QY 920 CATATCAGACGTACAGAGAGATTGAGAACAAACATG 958
DB 969 TTTATGTCACTACAGGAATGTTATGAGAAAATGATG 1007

RESULT 9
LOCUS OCSDBATRP 1116 bp mRNA linear MAM 12-OCT-1995
DEFINITION O. cuniculus mRNA for ileal sodium-dependent bile acid transporter.
ACCESSION 254357
VERSION 254357.1 GI:1019395
KEYWORDS ileal sodium-dependent bile acid transporter.
SOURCE Oryctolagus cuniculus.
ORGANISM Oryctolagus cuniculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1 (bases 1 to 1116)
AUTHORS Stengelin,S., Apel,S., Becker,W., Maier,W., Rosenberger,J., Wess,G.
and Kramer,W.
TITLE Cloning of the rabbit ileal sodium-dependent bile acid transporter
JOURNAL Unpublished
REFERENCES 2 (bases 1 to 1116)
AUTHORS Stengelin,S.
TITLE Direct Submision
JOURNAL Submitted (11-OCT-1995) Stengelin S., Hoechst Marion Roussel, TD
Metabolism, Building H825, D-65926 Frankfurt am Main, Germany
FEATURES
source location/Qualifiers
1..1116
/organism="Oryctolagus cuniculus"
/strain="New Zealand White"
/db_xref="taxon:9986"
/cruise_type="ileum"
/dev_stage="adult"
1..41
/partial
42..1085
/codon_start=1
/product="ileal sodium-dependent bile acid transporter"
/protein_id="CAA9184.1"
/db_xref="GI:1019396"
/db_xref="SWISS-PROT:Q28727"
/translation="MSNLTVGCLANATVGEASCVAPBSNFNAISVLSTVLTLLA
LWFSWGVNVEIKKFGHRIIRPWGIFIGFCQFGLMPLGFLYLAVFYGMPIQAVVL
IMGCCGATASNLIAWVDGMDLSVMTSCSTLALGMMPLCLVYGTMTWDSGRTV
IPYDNTGSLVALVYVSTVSGMFPNHNKPKAKITLKVSGIAGAVILVLAIVAGGILYQ
SMTIEPKMTITGITTPMAQYSLGFLALIAQGPWIKRKTVALETGMONTQICSTTVQ

3'UTR
BASE COUNT 265 a 274 c 283 g 294 t
ORIGIN
LSFSPEDITVFTPLPLYSIFQIAFMALFGLIYVANKKGNDAEPEDIKOTTEPE
SSRFQMGGFQPE"
1086..1116
/partial

Query Match 27.0%; Score 306.4; DB 4; Length 1116;
Best Local Similarity 59.2%; Pred. No. 1.6e-82;
Matches 523; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

QY 80 ATGGAACCTGAGCTGCTTTTCAAGTGTGTCCACTGTGATGATGGGCGCTCAAGT 139
DB 124 ATGCCATCTCAGCGCTGTGTTCTGAGTACCGGTGACCATCTGTGCTGTGATGT 183
QY 140 TCTCTTGGAGATGTTCCGTGAGATCCGGAAGCTGTGTCCACATCAGAGACCTGGG 199
DB 184 TCTCATGAGATGCACTGGAATCAAGAAATTCCTGGGCACTAAGCGGCGCTGG 243
QY 200 GCATTGCTGTGGAGCTGTGCGCAATTTGGAGCTTCATGCTTTTACAGCTTATCTCTGG 259
DB 244 GCATCTTCAATGGCTTCTCTGCGCAATTTGGAGATGATGCCCTCAGGGAATTTGCTTAG 303
QY 260 CCATTAGCTTTTCTGTGAAGCCAGTCCAGCTATTTGCTGTCTATCATATGGGCTGCTCC 319
DB 304 CGGTGCTTTTGGAGATGCTCCATTCAGCGCGGTGTGTCTCATATGAGGATGCTCC 363
QY 320 CGGGGGGACCATCTCTAATATTTTACCTTCTGGGTTGATGAGATGATGATCTCAGCA 379
DB 364 CTGAGAGAACGGCTCTCCAAATCTCTGTGATTTGGGTGATGAGACATGACCTTGAATG 423
QY 380 TCAGTATGACAACTGTGTCCACCGTGGCCCTGCGGAATGATGCTGCAATTTATC 439
DB 424 TCAGCATGACCACTGTCTCCACATTTGCTGCTGCAATGATGATGCTTATGCTTATG 483
QY 440 TCTACAGCTGTGCTGTGAGTCTTACAGAGATTCACCATTCCTTATGAGAAATAGGA 499
DB 484 TCTACACCAAAATGTTGGTGTGACCTGTGAGCCATTTATCTTACATACATATAGGA 543-
QY 500 TTACCTTGTGTGCTGTGACCAATCTCTGTGCGCTTGTGTGTATGATGATGATGATG 559
DB 544 CTCTCTGTGTGTCTGT 603
QY 560 CAAACCAATCCAAATCATTTCTCAAGATTGGGCGCTTGTGTGTGTGTGTGTGTGTGTGT 619
DB 604 CCAAAAGCGCAAGATTATCTTAAAGTTGATTCATTCAGGTGCGAGTCTTATTTGTGC 663
QY 620 TGGTCGAGT 679
DB 664 TCATGCTGT 723
QY 680 TGAACCATGATTTCACTTCTTCTTGTGATTTGGCCATGTACAGGCTTTTGTGTGTGTGT 739
DB 724 GGAATTATAGGAAGATTTTCCATGTGCGGTTATCTCCCTTGTGCTTTTGTGTGTGTGT 783
QY 740 TTACCCACAGCTTGTGCAAGGTGCGAGACATTTCTTGAATCTGAGTGTGAGATA 799
DB 784 TAGCTGTGTGAGCATGT 843
QY 800 TTCAGATGTGATGCACATGT 859
DB 844 CACAGCTGT 903
QY 860 TGAATTTCCCACTGTGCTATGAGACTCTTCCAGCTGATGATGATGATTTCTTATTTGTGAG 919
DB 904 TCACCTTCCCGCTCATTTTACAGCATCTTCCAGATGCGCTTTGACAGCAATCTTCTTAGGA 963
QY 920 CATATCAGACGTACAGAGAGATTGAAGAACAAACATGAGAAA 963
DB 964 TATATGTGCAATATGAGAAATGTCAATGAGAAAATGATGACAGA 1007

RESULT 10

AR033871 AR033871 1047 bp DNA linear PAT 29-SEP-1999
LOCUS Sequence 3 from patent US 5869265.
DEFINITION AR033871
ACCESSION AR033871
VERSION AR033871.1 GI:5949476
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1047)
Unclassified.
AUTHORS Dawson, P.A.
TITLE Ileal bile acid transporter compositions and methods
JOURNAL Patent: US 5869265-A 3 09-FEB-1999;
FEATURES Location/Qualifiers
source 1.1047
/organism="unknown"
BASE COUNT 251 a 251 c 255 g 290 t
ORIGIN
Query Match 26.3%; Score 297.8; DB 6; Length 1047;
Best Local Similarity 58.5%; Pred. No. 7.2e-80;
Matches 518; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 80 ATGGAACCTGAGCTCGTTTACAGTGTGCACTGTGATGATGGGCTGCTCATGT 139
DB 80 ATAACTCTTAAGTGTGCTTAAGTACGATGCTGATCCATCTGTTGGCTTGTGATGT 139
QY 140 TCTCTTTGGGATGTTCCGTGAGATCCGGAAGCTGTGTGCGACATCAGAGACCTGGG 199
DB 140 TCTCCATGGGATGCAACGTGGAATTCAGAAATTTCTAGGGCACTAAAGCCGCTGGG 199
QY 200 GCATTGCTGTGGAGCTGCTGCGCAGTTGGGGCTCATGCTTTACAGTTATCTCTGG 259
DB 200 GCATTGCTGTGGAGCTGCTGCGCAGTTGGGGCTCATGCTTTACAGTTATCTCTGG 259
QY 260 CCATTAGCTTTCTCTGAAGCCAGTCCAAAGCTATTGCTGTTCATCATGAGCTGTGCC 319
DB 260 CGGTGGCTTTGACATCTCCGCTCCAGCCGTAAGTGTGCTCATTAAGATGTGCTCC 319
QY 320 CGGGGGGCAACATCTCTAATTTTCACTTCTGGGTATGAGATATGATCTCAGCA 379
DB 320 CTGGAGGAACCTGCTCCAAATCTTGGCTATGAGTGTGATGCGACATGAGACCTGAGG 379
QY 380 TAGATATGCAACCTGTTCACCGTCCGCGCCCTGGGAATGATGCACTCTGATTTATC 439
DB 380 TAGATATGCAACCTGTTCACCGTCCGCGCCCTGGGAATGATGCACTCTGATTTATC 439
QY 440 TCTACACCTGTGCTGTGAGTCTTCAAGACATCTCAACATCTCTTATCAGAACATAGAA 499
DB 440 TCTATACCAAAATGTGGGTGCACTGTGGAGCATGTATTCCTATATATACATAGATA 499
QY 500 TTACCTGTGTGCTGACCAATTTCTGTGGCTTTGTGTGTATGATTAACAATGAGC 559
DB 500 CATCTGTGTGCTGCTGTGTGTCTGTTCCTATTTGGAATGTGTTATATCAAAATGAGC 559
QY 560 CAAAACATCCAAATCATTTCTCAAGATTGGGGCGTGTGGTGGGGCTCTCTCTGG 619
DB 560 CCAAAAAGCAAGATCATCTTAATTAATTTGGGTCAATGCGGGCGCATCTCATATGTGC 619
QY 620 TGTGCGAGTTGCTGTGTGTGCTGTGCGAAGATCTTGAATTCAGACATCACCTTC 679
DB 620 TATATGCTGTGTGAGGATATTTGATCAAAAGGCTGTGATCATGTCTCCAAATGT 679
QY 680 TACACATCATGTTCACTTTCTTTGATTTGSCCATGTCAAGGTTTCTGTGCGACATTT 739
DB 680 GATATTAAGAACAAATATTTCTGTGCGGTTACTCTCTGGGGTTTCTTGTGCGTAGAA 739
QY 740 TTACCAACAGTCTTGGCAAGGTGAGACAAATTTCTTAGAAATGAGGTGCGAATA 799
DB 740 TTGCTGTGTATCCCTGTGATGAGGTGCGAAGGTTCTTTTGAACGGGGATGCGAATA 799
QY 800 TTCAAGTGTGATCAGCATGCTCAAGTTATCTTTCACTGCTAGACACTTGTGTCAGATGT 859

DB 800 CGCAGTATGTTCCACATCTGTTACGCTCTCTTCACTCTGAGGAGCTCAATGTGTAT 859
QY 860 TGAATTTCCCACTGGCCATATGACTCTTCCAGCTCATATGATGATTTCTTATGTTGAG 919
DB 860 TCACCTTCCGCTCATCTACAGCATTTTCAGCTGCTGCTTGGCCGAATATTTTAGAT 919
QY 920 CATATCAGACGTACAGAGAGATTGAAGAACAAATGAGAAAA 964
DB 920 TTTATGTGATACAGAAATGTCAATGAAAAAACAAGCAGAAA 964
RESULT 11
LOCUS 132745 1047 bp DNA linear PAT 06-FEB-1997
DEFINITION Sequence 3 from patent US 5589358.
ACCESSION 132745
VERSION 132745.1 GI:1823536
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1047)
Unclassified.
AUTHORS Dawson, P.A.
TITLE Ileal bile acid transporter compositions and methods
JOURNAL Patent: US 5589358-A 3 31-DEC-1996;
FEATURES Location/Qualifiers
source 1.1047
/organism="unknown"
BASE COUNT 251 a 251 c 255 g 290 t
ORIGIN
Query Match 26.3%; Score 297.8; DB 6; Length 1047;
Best Local Similarity 58.5%; Pred. No. 7.2e-80;
Matches 518; Conservative 0; Mismatches 367; Indels 0; Gaps 0;
QY 80 ATGGAACCTGAGCTCGTTTACAGTGTGCACTGTGATGATGGGCTGCTCATGT 139
DB 80 ATAACTCTTAAGTGTGCTTAAGTACGATGCTGATCCATCTGTTGGCTTGTGATGT 139
QY 140 TCTCTTTGGGATGTTCCGTGAGATCCGGAAGCTGTGTGCGACATCAGAGACCTGGG 199
DB 140 TCTCCATGGGATGCAACGTGGAATTCAGAAATTTCTAGGGCACTAAAGCCGCTGGG 199
QY 200 GCATTGCTGTGGAGCTGCTGCGCAGTTGGGGCTCATGCTTTACAGTTATCTCTGG 259
DB 200 GCATTGCTGTGGAGCTGCTGCGCAGTTGGGGCTCATGCTTTACAGTTATCTCTGG 259
QY 260 CCATTAGCTTTCTCTGAAGCCAGTCCAAAGCTATTGCTGTTCATCATGAGCTGTGCC 319
DB 260 CGGTGGCTTTGACATCTCCGCTCCAGCCGTAAGTGTGCTCATTAAGATGTGCTCC 319
QY 320 CGGGGGGCAACATCTCTAATTTTCACTTCTGGGTATGAGATATGATCTCAGCA 379
DB 320 CTGGAGGAACCTGCTCCAAATCTTGGCTATGAGTGTGATGCGACATGAGACCTGAGG 379
QY 380 TAGATATGCAACCTGTTCACCGTCCGCGCCCTGGGAATGATGCACTCTGATTTATC 439
DB 380 TAGATATGCAACCTGTTCACCGTCCGCGCCCTGGGAATGATGCACTCTGATTTATC 439
QY 440 TCTACACCTGTGCTGTGAGTCTTCAAGACATCTCAACATCTCTTATCAGAACATAGAA 499
DB 440 TCTATACCAAAATGTGGGTGCACTGTGGAGCATGTATTCCTATATATACATAGATA 499
QY 500 TTACCTGTGTGCTGACCAATTTCTGTGGCTTTGTGTGTATGATTAACAATGAGC 559
DB 500 CATCTGTGTGCTGCTGTGTGTCTGTTCCTATTTGGAATGTGTTATATCAAAATGAGC 559
QY 560 CAAAACATCCAAATCATTTCTCAAGATTGGGGCGTGTGGTGGGGCTCTCTCTGG 619
DB 560 CCAAAAAGCAAGATCATCTTAATTAATTTGGGTCAATGCGGGCGCATCTCATATGTGC 619
QY 620 TGTGCGAGTTGCTGTGTGTGCTGTGCGAAGATCTTGAATTCAGACATCACCTTC 679

Db	620	TCATAGCTGTGGTGGAGAAATATGTACCAAGCGCTGGATCATTTGCTCCCAACTGT	679	/cissue_type="ileum"
Qy	680	TCACCATGATTTTCTTCTTTGATTTGGCCATGTCAGGGTTTCTGTGGCACTTT	739	/clone_id="human ileal cdna lambda gt10"
Db	680	GGATATATAGAACAAATATTTCTGTGGCGGTTACCTCGGGTTTCTTGGCTAGAA	739	/dev_stage="adult"
Qy	740	TTACCCACCAAGTCTTGGCAAGGTGACAGACAATTTCTTAGAAGTGGAGCTCAGATA	799	/gene="SLC10-A2"
Db	740	TTGCTGGTGTACCTGTGACAGGTGCGAAGCGGTCTTTTGAACGCGGATGCAAGACA	799	1..598
Qy	800	TTACAGATGTCATCAGCATGCTCAGTTATCTTCACTGTGAGCACTTGTCCAGATGT	859	/gene="SLC10-A2"
Db	800	CCGAGTATGTTCCACATGCTTTCACCTCTTCTTCACTGTGAGCACTTGTCCAGATGT	859	599..1645
Qy	860	TCAGTTTCCCATGCTGCTTATGACCTCTTCCAGCTGATAGATTTCTTATTTGTCAG	919	/gene="SLC10-A2"
Db	860	TCACCTTCCCGCTCATCTACAGATTTTCAGCTTCCCTTTGGCCGAAATATTTAGAT	919	/dev_stage="adult"
Qy	920	CATATACAGCTGTCAGAGAGATTGAAGAACAAATGAGAAA	964	/cissue_type="ileum"
Db	920	TTATGTGGCATACAGAAATGTATGAGAAAACAAAGCAGAAA	964	/clone_id="human ileal cdna lambda gt10"
RESULT 12	HSU10417	3779 bp mRNA linear PRI 02-OCT-2001		
LOCUS	HSU10417	3779 bp mRNA linear PRI 02-OCT-2001		
DEFINITION	Human ileal sodium-dependent bile acid transporter			
ACCESSION	U10417	GI:2623285		
VERSION	U10417.1	GI:2623285		
KEYWORDS	Human sapiens			
SOURCE	Human sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 3779)			
AUTHORS	Wong, M.H., Oelkers, P. and Dawson, P.A.			
TITLE	Identification of a mutation in the ileal sodium-dependent bile acid transporter gene that abolishes transporter activity			
JOURNAL	J. Biol. Chem. 270 (45), 27228-27234 (1995)			
MEDLINE	96070831			
PUBMED	7592981			
REFERENCE	2 (bases 1 to 3779)			
AUTHORS	Craddock, A.L., Love, M.W., Daniel, R.W., Kirby, L.C., Walters, H.C., Wong, M.H. and Dawson, P.A.			
TITLE	Expression and transport properties of the human ileal and renal sodium-dependent bile acid transporter			
JOURNAL	Am. J. Physiol. 274 (1 Pt 1), G157-G169 (1998)			
MEDLINE	98120406			
PUBMED	9458785			
REFERENCE	3 (bases 1 to 3779)			
AUTHORS	Dawson, P.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (07-JUN-1994) Paul A. Dawson, Internal Medicine/Gastroenterology, Bowman Gray School of Medicine, Wake Forest University, Medical Center Boulevard, Winston-Salem, NC 27157, USA			
REFERENCE	4 (bases 1 to 3779)			
AUTHORS	Dawson, P.A.			
TITLE	Direct Submission			
JOURNAL	Submitted (17-NOV-1997) Paul A. Dawson, Internal Medicine/Gastroenterology, Bowman Gray School of Medicine, Wake Forest University, Medical Center Boulevard, Winston-Salem, NC 27157, USA			
REMARK	Sequence update by submitter			
COMMENT	On Nov 17, 1997 this sequence version replaced gi:595398.			
FEATURES	Location/Qualifiers			
source	1..3779			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/map="13g33"			
	/clone="PHISBT"			

Qy	80	ATGAAACCTGGAGCTGCTTTCACAGTGTGTCACATGATGATGAGGCTGCTCATGT	139	/cissue_type="ileum"
Db	678	ATACATCTTAAGTGTGCTTCCATGATGATGATGATGATGATGATGATGATGATGAT	737	/clone_id="human ileal cdna lambda gt10"
Qy	140	TCTCTTTGGAGTGTTCCTGAGATCCGGAAGCTGTGTCACATCAGAGACCTCTGG	199	/dev_stage="adult"
Db	738	TCTCCATGAGATGACCACTGAAATCAAAATTTCTAGGCACTAAAGCGCTGTGG	797	/gene="SLC10-A2"
Qy	200	GCATGCTGTGGAGTGTGCTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	259	1..598
Db	798	GCATGCTGTGGAGTGTGCTGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	857	/gene="SLC10-A2"
Qy	260	CCATTAGCTTCTCTGCAAGCGAGTATGCTGTCTCATCATGAGGCTGTGCTGCTG	319	599..1645
Db	858	CGGTGGCTTTGACATCTCTCCGCTCCAGGCGGTGTGTGTGTGTGTGTGTGTGTGT	917	/gene="SLC10-A2"
Qy	320	CGGGGGGACCATCTCTTAACATTTTACCTTGTGCTGTGATGATGATGATGATGATG	379	/dev_stage="adult"
Db	918	CTGAGGAACTGCTCCCAATATCTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	977	/cissue_type="ileum"
Qy	380	TCAGTATGCAACTGTTCACCGTGGCGGCTGGAATGATGATGATGATGATGATGATG	439	/clone_id="human ileal cdna lambda gt10"
Db	978	TCAGTATGCAACTGTTCACCGTGGCGGCTGGAATGATGATGATGATGATGATGATG	1037	/dev_stage="adult"
Qy	440	TCTACACCTGTGCTGAGTCTTTCAGCAATCTGCAATCTTATGCAAAATAGAA	499	/gene="SLC10-A2"
Db	1038	TCTATACCAAAATGAGGAGCACTGTGAGAGATGATGATGATGATGATGATGATG	1097	1..598
Qy	500	TTACCTTGTGCTGCAACTTCTGTGCTGCTTGTGCTGTGCTGTGCTGTGCTGTG	559	/gene="SLC10-A2"
Db	1098	CATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1157	599..1645
Qy	560	CAAAACATCCAAATATCTCAAGATTTGGGGCGTGTGCTGCTGCTGCTGCTGCTG	619	/gene="SLC10-A2"
Db	1158	CCCAAAACATCCAAATATCTCAAGATTTGGGGCGTGTGCTGCTGCTGCTGCTG	1217	/dev_stage="adult"
Qy	620	TGCTGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	679	/cissue_type="ileum"
Db	1218	TCATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1277	/clone_id="human ileal cdna lambda gt10"
Qy	680	TCACCATGATTTTCTTCTTTGATTTGGCCATGTCAGGGTTTCTGTGGCACTTT	739	/dev_stage="adult"
Db	1278	GGATATATAGAACAAATATTTCTGTGGCGGTTACCTCGGGTTTCTTGGCTAGAA	1337	/gene="SLC10-A2"
Qy	740	TTACCCACCAAGTCTTGGCAAGGTGACAGACAATTTCTTAGAAGTGGAGCTCAGATA	799	1..598

Db	1338	TTGCTGGTGTCTACCCCTGGATKAGAGTCGCCAAGCGTTGCTTTTGAAACGGSGATGCAGAACA	1337
Oy	800	TTCAGAFGTGCATCACCATGCTCCAGTTATCTTTCACTGCTGAGCACTTGTCCAGATGT	859
Db	1398	CGCAGACTATGTTCCACCACATCGTTGACGCTCTCCTTCACCTCGAGGAGCTCAATGCGTAT	1457
Oy	860	TGAGTTTCCCACCTGGCCATTATGAGACCTCTCCAGCTGATATGAATGATTTCTTAATGTGAC	919
Db	1458	TCACCTTCCCCGTCACTTCACAGCATTTTTCAGCTGCGCTTTGCCGAATATTTCTTAGAT	1517
Oy	920	CATATCAGACGTCACAGAGAGATTTGAAGAACCAATCATGAAAAA	964
Db	1518	TTTATGTGCATACCAAGAAATGTATGTAGAAAAAACAGCGAGAA	1562
RESULT 13			
LOCUS	RNU07183		
DEFINITION	RNU07183	4269 bp	mRNA linear ROD 16-JUL-2001
ACCESSION	RNU07183		
VERSION	U07183		
KEYWORDS	U07183.2 GI:14719390		
SOURCE			
ORGANISM	Rattus norvegicus.		
REFERENCE	Rattus norvegicus.		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. 1 (bases 1 to 4269) Shneider,B.L., Dawson,P.A., Christie,D.M., Hardikar,W., Wong,M.H. and Suchy,F.J. Cloning and molecular characterization of the ontogeny of a rat ileal sodium-dependent bile acid transporter J. Clin. Invest. 95 (2), 745-754 (1995) 95164708 7860756 2 (bases 1 to 4269) Christie,D.M., Dawson,P.A., Thevananthar,S. and Shneider,B.L. Comparative analysis of the ontogeny of a sodium-dependent bile acid transporter in rat kidney and ileum Am. J. Physiol. 271 (2 Pt 1), G377-G385 (1996) 96365877 8770054 3 (bases 1 to 4269) Chen,F., Ma,L. and Shneider,B.L. Transcriptional regulation of the rat apical sodium-dependent bile acid transporter: Basic mechanisms and response to inflammation and corticosteroids Unpublished 4 (bases 1 to 4269) Shneider,B.L. Direct Submission Submitted (25-FEB-1994) Benjamin L. Shneider, Pediatrics, Yale University, 333 Cedar Street, New Haven, CT 06520, USA 5 (bases 1 to 4269) Shneider,B.L. Direct Submission Submitted (11-JUN-2000) Pediatrics, Mount Sinai School of Medicine, One Gustave L. Levy Place (Box 1656), NY, NY 10029, USA Sequence update by submitter On Jul 16, 2001 this sequence version replaced gi:1488241. Location/Qualifiers 1..4269 /organism="Rattus norvegicus" /strain="Sprague-Dawley" /db_xref="taxon:10116" /clone="BS37C1" /tissue_type="Ileum" /dev_stage="adult" 1..107 108..1154 /codon_start=1 /product="Ileal sodium-dependent bile acid transporter"		

[illegible]

QY 949 AACAAACATG 958
 Db 1056 AAAAATGATG 1065

RESULT 14
 AC120684
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-13M5. *** SEQUENCING IN PROGRESS ***
 52 unordered pieces.
 AC120684
 AC120684.2 GI:21902828
 HTG: HTGS_Phasel.
 KEYWORDS Rattus norvegicus.
 SOURCE Rattus norvegicus.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 106811)
 AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T., Barabasi,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Cartron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthett,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homai,P., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kretovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,D., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Louieged,H., Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E., Massey,B., Maunney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabhat,K., Morgan,M., Morris,S., Moser,M., Newtson,J., Newton,S.N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okunolu,G., Odegunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Pimus,E., Pul,L., Quiles,M., Ren,Y., Rives,M., Rojase,A., Rojuben,I., Rolle,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I., Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H., Suton,A., Svatek,A., Tabors,P., Tamerisa,A., Tamerisa,K., Tang,H., Taney,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umani,K., Vaquer,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G. and Gibbs,R.

TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 106811)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (09-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE 3 (bases 1 to 106811)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (23-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Jul 18, 2002 this sequence version replaced gi:20514546.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information

Center project name: GANE
 Center clone name: CH230-13M5
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 62816 bases at least Q40
 Consensus quality: 68678 bases at least Q30
 Consensus quality: 72899 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 52 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1673: contig of 1673 bp in length
 1 1674 1773: gap of unknown length
 1 1774 3396: contig of 1623 bp in length
 1 3397 3496: gap of unknown length
 1 3497 4846: contig of 1350 bp in length
 1 4847 4947 6033: contig of 1087 bp in length
 1 4947 6034 7730: gap of unknown length
 1 6034 7730: contig of 1597 bp in length
 1 7731 7830: gap of unknown length
 1 7831 9114: contig of 1284 bp in length
 1 9115 10568: contig of 1354 bp in length
 1 9215 10568: gap of unknown length
 1 10569 11701: contig of 1033 bp in length
 1 10669 11702 11801: gap of unknown length
 1 11802 13342: contig of 1541 bp in length
 1 13343 13442: gap of unknown length
 1 13443 14661: contig of 1219 bp in length
 1 14662 14761: gap of unknown length
 1 14762 16156: contig of 1395 bp in length
 1 16157 16256: gap of unknown length
 1 16257 17467: contig of 1211 bp in length
 1 17468 17567: gap of unknown length
 1 17568 18876: contig of 1309 bp in length
 1 18877 18976: gap of unknown length
 1 18977 21088: contig of 2112 bp in length
 1 21089 21188: gap of unknown length
 1 21189 22267: contig of 1079 bp in length
 1 22268 22367: gap of unknown length
 1 22368 23405: contig of 1038 bp in length
 1 23406 23505: gap of unknown length
 1 23506 24559: contig of 1054 bp in length
 1 24560 24659: gap of unknown length
 1 24660 26033: contig of 1374 bp in length
 1 26034 26133: gap of unknown length
 1 26134 27715: contig of 1582 bp in length
 1 27716 27815: gap of unknown length
 1 27816 29020: contig of 1205 bp in length
 1 29021 29120: gap of unknown length
 1 29121 30180: contig of 1060 bp in length
 1 30181 30280: gap of unknown length
 1 30281 31631: contig of 1351 bp in length
 1 31632 31731: gap of unknown length
 1 31732 33324: contig of 1593 bp in length
 1 33325 33424: gap of unknown length


```

* 33425 34541: contig of 1117 bp in length
* 34542 34541: gap of unknown length
* 34642 36373: contig of 1732 bp in length
* 36374 36473: gap of unknown length
* 36474 37986: contig of 1513 bp in length
* 37987 38086: gap of unknown length
* 38087 40644: contig of 2558 bp in length
* 40645 40744: gap of unknown length
* 40745 42886: contig of 2142 bp in length
* 42887 42986: gap of unknown length
* 42987 45107: contig of 2121 bp in length
* 45108 45207: gap of unknown length
* 45208 47360: contig of 2153 bp in length
* 47361 47460: gap of unknown length
* 47461 49278: contig of 1718 bp in length
* 49279 50397: contig of 1119 bp in length
* 50398 50497: gap of unknown length
* 50498 52192: contig of 1695 bp in length
* 52193 52292: gap of unknown length
* 52293 54790: contig of 2498 bp in length
* 54791 54890: gap of unknown length
* 54891 56793: contig of 1903 bp in length
* 56794 56893: gap of unknown length
* 56894 59215: contig of 2322 bp in length
* 59216 59315: gap of unknown length
* 59316 62155: contig of 2840 bp in length
* 62156 62255: gap of unknown length
* 62256 64324: contig of 2069 bp in length
* 64325 64424: gap of unknown length
* 64425 67073: contig of 2649 bp in length
* 67074 67173: gap of unknown length
* 67174 70259: contig of 3086 bp in length
* 70260 70359: gap of unknown length
* 70360 73801: contig of 3442 bp in length
* 73802 73901: gap of unknown length
* 73902 76070: contig of 2169 bp in length
* 76071 76170: gap of unknown length
* 76171 78275: contig of 2105 bp in length
* 78276 78375: gap of unknown length
* 78376 81217: contig of 2842 bp in length
* 81218 81317: gap of unknown length
* 81318 83901: contig of 2584 bp in length
* 83902 84001: gap of unknown length
* 84002 87212: contig of 3211 bp in length
* 87213 87312: gap of unknown length
* 87313 89028: contig of 1716 bp in length
* 89029 89128: gap of unknown length
* 89129 91539: contig of 2411 bp in length
* 91540 91639: gap of unknown length
* 91640 95692: contig of 4053 bp in length
* 95693 95792: gap of unknown length
* 95793 99538: contig of 3746 bp in length
* 99539 99638: gap of unknown length
* 99639 102477: contig of 2839 bp in length
* 102478 102577: gap of unknown length
* 102578 106811: contig of 4224 bp in length.

```

```

FEATURES
    source
        1..106811
            /organism="Rattus norvegicus"
            /db_xref="taxon:10116"

```

```

Query Match      23.4%; Score 265; DB 2; Length 106811;
Best Local Similarity 81.4%; Pted. No. 1.9e-69;
Matches 307; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```

```

Qy 1 ATGAGAGCAATTGTTCCAGCAGCTGCTCCCTGCAACAGTTCAAGAGAGAGCTG 60
Db 55006 ATGAGCGCAGACTCGGAGGCGCACTCCCTGCGCAACAGCAGGAGAAACCCA 55065
Qy 61 CCAAGTGGAGCTGGAGGTGCATGAAACCTGAGAGCTGTTTCAAGTGTGTCTGCTGTG 120
Db 55066 CCGGTGGAGATGAGAGGACAGGGGAGCCCTGAGCTGTTTTCACAGTCTGTGGCTGTG 55125

```

```

Qy 121 ATGATGGGGCTGCTCATGTTCTTTTGGGATGTTCCGAGAGATCCGGAAGCTGNGTCG 180
Db 55126 ATGATGGGTCTGCTCATGTTCTTTTGGGATGTTCCGAGAGATCCGGAAGCTGNGTCG 55185
Qy 181 CACATCAGAGAGACCCCTGGGAGATTGCTGTGAGCTGCTCCAGTTTGGCTCAGTCCCT 240
Db 55186 CACCTCAGAAACCCCTGGGAGATGAGAGTGGGCGCTTTCAGATTGGGCTCAGTCCCT 55245
Qy 241 TTTACAGCTTATCTCTGCGCATTAAGCTTTTCTGAAAGCAGATCAGATTTGCTGTT 300
Db 55246 CTGACAGCTTATCTCTGCGCATTAAGCTTTTCTGAAAGCAGATCAGATTTGCTGTT 55305
Qy 301 CTGATCATGGAGCTGCGCGGGGGGACACATCTCTAATTTTACCTTCGAGTTGAT 360
Db 55306 CTGATCATGGAGAGTCTGCGGGGACCGCTCTTATATGCTCATCCTTGGGTGAT 55365
Qy 361 GGAGATATGATCTCAG 377
Db 55366 GGAGATATGATCTCAG 55382

```

```

RESULT 15
AL713989 215210 bp DNA linear ROD 16-AUG-2002
LOCUS Mouse DNA sequence from clone RP23-514 on chromosome 5, complete
DEFINITION
ACCESSION AL713989
VERSION AL713989
KEYWORDS GI:22316185
SOURCE HTG.
ORGANISM Mus musculus.
            Mus musculus.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
            1 (bases 1 to 215210)
            Direct Submission
            Submitted (16-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
            On Aug 19, 2002 this sequence version replaced gi:21272445.
            ----- Genome Center
            Center: UK Medical Research Council
            Center code: UK-MRC
            Web site: http://mrcseq.har.mrc.ac.uk
            Contact: mouseq@har.mrc.ac.uk

```

COMMENT

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission. Only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, EMBL; Swi, SWISSPROT; Tr, TrEMBL; Mp, WormPep; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-514 is from the RPI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6.

```

FEATURES
    source
        1..215210
            /organism="Mus musculus"
            /db_xref="taxon:10090"

```

```

/chromosome="5"
/clone="RP23-514"
/clone_lib="RPc1-23"
BASE COUNT      59453 a 48517 c 47922 g 59318 t
ORIGIN

```

```
/clone="RP23-514"
```

a 48517 c 47922 g

BASE COUNT	59453	a	48517	c	47922	g	59318	e
ORIGIN								

ORIGIN

Query Match	23.1%	Score 261.8;	DB 10;	Length 215210;
Best Local Similarity	80.9%;	Pred. No. 2e-68;		
Matches 305;	Conservative 0;	Mismatches 72;	Indels 0;	Gaps 0;

Best Local Similarity 80.9%; Pred. No. 2e-68;

Matches 305; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

Oy	1	ATGAGCCCAATTGTTCTTCACGACGCTCAGCTGCTGCCCGCAACAGTTCTCAAGGAGGAGCTG	60
Db	139053	ATGAGACACAGACTGTGTCGGGCACTCCACTGCTCCCTGTCAACAGTACGAGGAGGAGACCG	139112
Oy	61	CCAGTGGGACTTGAGGTGCATGGAAACCTGAGAGCTGGTTTCA CAGTGGTGTCACTGTG	120
Db	139113	CCCGTGGGAATGAGAGGGCCATGGGAATCTAAAGCTCCTTTTTCAGTGTCTCTGGCTGTG	139172
Oy	121	ATGATGGGGCTGCTCATGTTCTCTTTGGGATGTTCCGTGGAGATCCGGAAGCTGTGCTG	180
Db	139173	ATGGTGGGTTTGTCATGTTCTCTTTTGGATGTTCTGTGGAGAGTCAGAAAGCTCTGGTTG	139232
Oy	181	CACATCAGGAGACCCCTGGGGGCAATTGCTGTGGGACGTCTCTGACAGTTTGGGCTCATGCC	240
Db	139223	CACCTCAGAGAGCCCTTGGGGCATCGAGTGGGCTGCTCTTCCAGTTTGGACTTAATGCC	139292
Oy	241	TTTACAGCTTATCTCTCGGCACATTAGCTTTTCTGTGAGCCAGTCCAGCTATTGCTGT	300
Db	139293	CTGACAGCTTATCTGTATTGACATTGGCTTCGGTCTGAACCACTTCCAGGTATTGCTGTG	139352
Oy	301	CTCATCATGGGCTGCTGCCCGGGGGGACCATCTTAAACATTTTCACTTTCTGGTTGAT	360
Db	139353	CTCATGATGGGAGGCTGCCCTGGGGGACACATCTTAAAGTTCTCACTTCTGGTTGAT	139412
Oy	361	GGAGATATGATCTCAG	377
Db	139413	GGAGATATGATCTCAG	139429

Db 139053 ATGAGCACAGACTGTGCGGCACTCCACCTGCCCTGTCAACAGTACGGAGAGACCCG 139112

61 CCAGTGGGACTGGAGGTCATGGAACCTGGAGCTCGTTTTCACAGTGGTGTCCACTGTG 120

Db 139113 CCGTGGGATGGAGGGCCATGCCAATCTAAAGCTGCTTTTACAGTGTCTCGGCTGTG 139172

121 ATGATGGGCTGCTCATGTTCTTTGGGATGTTCCGTGGAGATCCGGAGCTGTGTCG 180

139173 ATGGTGGGTTTGGTCATGTTCTTTTGGATGTTCTGTGGAGAGTCAGAAGCTCTGGTTG 139232

181 CACATCAGGAGACCTGGGCAATGCTGAGACTGCTCTGCGAGTTGGCTCAGGCT 240

Db 139233 CACCTCAGAAGACCCCTGGGCATCGCAGTGGGCCCTTCCAGTTGGACTTATGCT 139292

241 TTACAGCTTATCTCCTGGCCATTAGCTTTTCTCTGAAGCCAGTCCAGCTATMGCTGT 300

Db 139293 CTGACAGCTATCTGTTAGCCATTGGCTTCGGTCTGAACCATTTCCAAGCTATTGCTGTC 139352

301 CTCAATCAATGGGCTGCTGCCCCGGGGCACCAATCTCTAAATTTTCACCTTCTGGGTGAT 360

Db 139353 CTTCAATGATGGGAGCTGGCCCTGGGGGCAACCATCTCTAATGTTCTCACCCTCTGGGTTGAT 139412

361 GGAGATATGGATCTCAG 377

Db ~ 139413 GGAGATATGGATCTCAG 139429

Search completed: June 9, 2003, 06:27:26
Job time : 3078 secs

Job time : 3078 secs